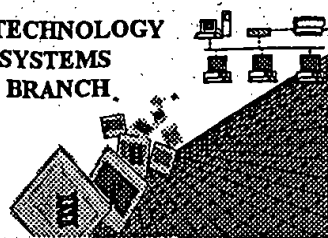


T. Bhatti

Re-run

**RAW SEQUENCE LISTING**  
**ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/636,243  
Source: 1600  
Date Processed by STIC: 2/21/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 09/636,243

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☐ Wrapped Nucleics  
Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to 3; this will prevent "wrapping."
- 2 ☐ Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ☐ Misaligned Amino Numbering The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 ☐ Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ☒ Variable Length Sequence(s) 111 contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 ☐ PatentIn 2.0 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)           . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 ☐ Skipped Sequences (OLD RULES) Sequence(s)            missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 ☐ Skipped Sequences (NEW RULES) Sequence(s)            missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9 ☐ Use of n's or Xaa's (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ☐ Invalid <213> Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ☐ Use of <220> Sequence(s)            missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 ☐ PatentIn 2.0 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 ☐ Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,243

DATE: 02/21/2002

TIME: 11:29:22

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\02212002\I636243.raw

Does Not Comply  
Corrected Diskette Needed

pp 1, 3-4

3 <110> APPLICANT: WANG, Bryan S.  
4 PABO, Carl O.  
6 <120> TITLE OF INVENTION: DIMERIZING PEPTIDES  
8 <130> FILE REFERENCE: 8325-1004 / M4-US1  
10 <140> CURRENT APPLICATION NUMBER: 09/636,243  
11 <141> CURRENT FILING DATE: 2000-08-10  
13 <150> PRIOR APPLICATION NUMBER: 60/148,422  
14 <151> PRIOR FILING DATE: 1999-08-11  
16 <160> NUMBER OF SEQ ID NOS: 83  
18 <170> SOFTWARE: PatentIn Ver. 2.0  
20 <210> SEQ ID NO: 1  
21 <211> LENGTH: 18  
22 <212> TYPE: PRT  
23 <213> ORGANISM: Artificial Sequence  
25 <220> FEATURE:  
26 <223> OTHER INFORMATION: Description of Artificial Sequence: exemplary  
27 motif characterizing C2H2 class proteins  
29 <220> FEATURE:  
30 <221> NAME/KEY: REPEAT  
31 <222> LOCATION: (2)  
32 <223> OTHER INFORMATION: where 2-4 Xaa's are present  
34 <220> FEATURE:  
35 <221> NAME/KEY: REPEAT  
36 <222> LOCATION: (17)  
37 <223> OTHER INFORMATION: where 3-5 Xaa's are present  
39 <220> FEATURE:  
40 <221> NAME/KEY: SITE  
41 <222> LOCATION: (2)  
42 <223> OTHER INFORMATION: where Xaa is any amino acid  
44 <220> FEATURE:  
45 <221> NAME/KEY: SITE  
46 <222> LOCATION: (4)..(15)  
47 <223> OTHER INFORMATION: where Xaa is any amino acid  
49 <220> FEATURE:  
50 <221> NAME/KEY: SITE  
51 <222> LOCATION: (17)  
52 <223> OTHER INFORMATION: where Xaa is any amino acid  
54 <400> SEQUENCE: 1  
W--> 55 Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His  
56 1 5 10 15  
W--> 58 Xaa His  
62 <210> SEQ ID NO: 2  
63 <211> LENGTH: 4

invalid - see item 5  
on Eva Summary  
sheet

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,243

DATE: 02/21/2002

TIME: 11:29:22

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\02212002\I636243.raw

```

64 <212> TYPE: PRT
65 <213> ORGANISM: Artificial Sequence
67 <220> FEATURE:
68 <223> OTHER INFORMATION: Description of Artificial Sequence: D-able
69     subsite
71 <400> SEQUENCE: 2
72 Asn Asn Gly Lys
73     1
76 <210> SEQ ID NO: 3
77 <211> LENGTH: 9
78 <212> TYPE: DNA
79 <213> ORGANISM: Artificial Sequence
81 <220> FEATURE:
82 <223> OTHER INFORMATION: Description of Artificial Sequence: zinc finger
83     protein bind sequence
85 <400> SEQUENCE: 3
86 ggcgtagac                                     9
88 <210> SEQ ID NO: 4
89 <211> LENGTH: 9
90 <212> TYPE: DNA
91 <213> ORGANISM: Artificial Sequence
93 <220> FEATURE:
94 <223> OTHER INFORMATION: Description of Artificial Sequence: zinc finger
95     protein bind sequence
97 <400> SEQUENCE: 4
98 ggcgacgta                                     9
100 <210> SEQ ID NO: 5
101 <211> LENGTH: 5
102 <212> TYPE: PRT
103 <213> ORGANISM: Artificial Sequence
105 <220> FEATURE:
106 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide
107     linker
109 <400> SEQUENCE: 5
110 Thr Gly Glu Lys Pro
111     1             5
114 <210> SEQ ID NO: 6
115 <211> LENGTH: 5
116 <212> TYPE: PRT
117 <213> ORGANISM: Artificial Sequence
119 <220> FEATURE:
120 <223> OTHER INFORMATION: Description of Artificial Sequence: linker
122 <400> SEQUENCE: 6
123 Gly Gly Gly Gly Ser
124     1             5
127 <210> SEQ ID NO: 7
128 <211> LENGTH: 8
129 <212> TYPE: PRT
130 <213> ORGANISM: Artificial Sequence

```

## RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/636,243

TIME: 11:29:22

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\02212002\I636243.raw

132 <220> FEATURE:  
133 <223> OTHER INFORMATION: Description of Artificial Sequence: linker  
135 <400> SEQUENCE: 7  
136 Gly Gly Arg Arg Gly Gly Ser  
137 1 5  
140 <210> SEQ ID NO: 8  
141 <211> LENGTH: 9  
142 <212> TYPE: PRT  
143 <213> ORGANISM: Artificial Sequence  
145 <220> FEATURE:  
146 <223> OTHER INFORMATION: Description of Artificial Sequence: linker  
148 <400> SEQUENCE: 8  
149 Leu Arg Gln Arg Asp Gly Glu Arg Pro  
150 1 5  
153 <210> SEQ ID NO: 9  
154 <211> LENGTH: 12  
155 <212> TYPE: PRT  
156 <213> ORGANISM: Artificial Sequence  
158 <220> FEATURE:  
159 <223> OTHER INFORMATION: Description of Artificial Sequence: linker  
161 <400> SEQUENCE: 9  
162 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro  
163 1 5 10  
166 <210> SEQ ID NO: 10  
167 <211> LENGTH: 16  
168 <212> TYPE: PRT  
169 <213> ORGANISM: Artificial Sequence  
171 <220> FEATURE:  
172 <223> OTHER INFORMATION: Description of Artificial Sequence: linker  
174 <400> SEQUENCE: 10  
175 Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro  
176 1 5 10 15  
179 <210> SEQ ID NO: 11  
180 <211> LENGTH: 18  
181 <212> TYPE: PRT  
182 <213> ORGANISM: Artificial Sequence  
184 <220> FEATURE:  
185 <223> OTHER INFORMATION: Description of Artificial Sequence: component  
186 finger of zinc finger protein  
188 <220> FEATURE:  
189 <221> NAME/KEY: REPEAT  
190 <222> LOCATION: (2)  
191 <223> OTHER INFORMATION: where 2-4 Xaa's are present  
193 <220> FEATURE:  
194 <221> NAME/KEY: REPEAT  
195 <222> LOCATION: (17)  
196 <223> OTHER INFORMATION: where 3-5 Xaa's are present  
198 <220> FEATURE:  
199 <221> NAME/KEY: SITE

*see item 5*

*item 5*

## RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/636,243

TIME: 11:29:22

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\02212002\I636243.raw

```

200 <222> LOCATION: (2)
201 <223> OTHER INFORMATION: where Xaa is any amino acid
203 <220> FEATURE:
204 <221> NAME/KEY: SITE
205 <222> LOCATION: (4)..(15)
206 <223> OTHER INFORMATION: where Xaa is any amino acid
208 <220> FEATURE:
209 <221> NAME/KEY: SITE
210 <222> LOCATION: (17)
211 <223> OTHER INFORMATION: where Xaa is any amino acid
213 <400> SEQUENCE: 11
W--> 214 Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His
215      1      5      10      15
W--> 217 Xaa His
221 <210> SEQ ID NO: 12
222 <211> LENGTH: 30
223 <212> TYPE: PRT
224 <213> ORGANISM: Artificial Sequence
226 <220> FEATURE:
227 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
228     domain F1
230 <400> SEQUENCE: 12
231 Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp
232  1      5      10      15
234 Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
235      20      25      30
238 <210> SEQ ID NO: 13
239 <211> LENGTH: 28
240 <212> TYPE: PRT
241 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
245     domain F2
247 <400> SEQUENCE: 13
248 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu
249  1      5      10      15
251 Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro
252      20      25
255 <210> SEQ ID NO: 14
256 <211> LENGTH: 38
257 <212> TYPE: DNA
258 <213> ORGANISM: Artificial Sequence
260 <220> FEATURE:
261 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
262     site
264 <400> SEQUENCE: 14
265 ggttgccagtg ggcgcgccca cagtacttga acgtaacg
266      38
267 <210> SEQ ID NO: 15
268 <211> LENGTH: 34

```

## RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/636,243

TIME: 11:29:22

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\02212002\I636243.raw

```

269 <212> TYPE: DNA
270 <213> ORGANISM: Artificial Sequence
272 <220> FEATURE:
273 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
274     site
276 <400> SEQUENCE: 15
277 cgttacgttc aagtactgtg ggcgcgccca ctgc 34
279 <210> SEQ ID NO: 16
280 <211> LENGTH: 12
281 <212> TYPE: DNA
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
286     site
288 <400> SEQUENCE: 16
289 tgggcgtatg ct 12
291 <210> SEQ ID NO: 17
292 <211> LENGTH: 12
293 <212> TYPE: DNA
294 <213> ORGANISM: Artificial Sequence
296 <220> FEATURE:
297 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
298     site
300 <400> SEQUENCE: 17
301 agcatacgcc ca 12
303 <210> SEQ ID NO: 18
304 <211> LENGTH: 57
305 <212> TYPE: DNA
306 <213> ORGANISM: Artificial Sequence
308 <220> FEATURE:
309 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
310     site
312 <400> SEQUENCE: 18
313 ggaattcctg atcaagatct ggtaacgtcc ataggctagg catgtcaagg ctgtatg 57
315 <210> SEQ ID NO: 19
316 <211> LENGTH: 57
317 <212> TYPE: DNA
318 <213> ORGANISM: Artificial Sequence
320 <220> FEATURE:
321 <223> OTHER INFORMATION: Description of Artificial Sequence: DNA binding
322     site
324 <400> SEQUENCE: 19
325 gggatccact cgcgaacgcg tcctttagt gggcgcgcc acatacagcc ttgacat 57
327 <210> SEQ ID NO: 20
328 <211> LENGTH: 12
329 <212> TYPE: DNA
330 <213> ORGANISM: Artificial Sequence
332 <220> FEATURE:
333 <223> OTHER INFORMATION: Description of Artificial Sequence: inverted

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/636,243

DATE: 02/21/2002

TIME: 11:29:23

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\02212002\I636243.raw

L:55 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:58 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:214 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11  
L:217 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11